

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Hadlaczky, Gyula
Szalay, Aladar
- (ii) TITLE OF THE INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF
AND METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Fish & Richardson
 - (B) STREET: 12390 El Camino Real
 - (C) CITY: San Diego
 - (D) STATE: CA
 - (E) COUNTRY: USA
 - (F) ZIP: 92130
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ Version 1.5
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: Herewith
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Seidman, Stephanie L.
 - (B) REGISTRATION NUMBER: 33,779
 - (C) REFERENCE/DOCKET NUMBER: 24601-4020
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 858-678-4777
 - (B) TELEFAX: 202-626-7796
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 1293 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: Genomic DNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTISENSE: NO
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
- (ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GAATTCATCA	TTTTTCANGT	CCTCAAGTGG	ATGTTTCTCA	TTNCCATGA	TTTTAAGTTT	60
TCTCGCCATA	TTCTGGTCC	TACAGTGTGC	ATTTCTCCAT	TTNCACGTT	TTNCAGTGAT	120
TTCGTCATTT	TCAAGTCCTC	AAGTGGATGT	TTCTCATTTN	CCATGAATTT	CAGTTTTCTN	180
GCCATATTCC	ACGTCTACA	GNGGACATTT	CTAAATTTC	CACCTTTTTC	AGTTTTCCCTC	240

GCCATATTTT	ACGTCCTAAA	ATGTGTATTT	CTCGTTTNCC	GTGATTTTCA	GTTTTCTCGC	300
CAGATTCCAG	GTCCTATAAT	GTGCATTTCT	CATTTNNCAC	GTTTTTCAGT	GATTTTCGTCA	360
TTTTTTCAAG	TCGGCAAGTG	GATGTTTCTC	ATTTNCCATG	ATTTNCAGTT	TTCTTGNAAT	420
ATTCCATGTC	CTACAATGAT	CATTTTTAAT	TTTCCACCTT	TTTCAATTTT	CACGCCATAT	480
TTTCATGTCCT	AAAGTGTATA	TTTCTCCTTT	TCCGCGATTT	TCAGTTTTCT	CGCCATATTC	540
CAGGTCCTAC	AGTGTGCATT	CCTCATTTTT	CACCTTTTTT	ACTGATTTCG	TCATTTTTTCA	600
AGTCGTCAAC	TGGATCTTTC	TAATTTTCCA	TGATTTTCAG	TTATCTTGTC	ATATTCCATG	660
TCCTACAGTG	GACATTTCTA	AATTTTCCAA	CTTTTTCAAT	TTTTCTCGAC	ATATTTGACG	720
TGCTAAAGTG	TGTATTTCTT	ATTTTCCGTG	ATTTTCAGTT	TTCTCGCCAT	ATTCCAGGTC	780
CTAATAGTGT	GCATTTCTCA	TTTTTTCACGT	TTTTTCAGTGA	TTTCGTCATT	TTTTCCAGTT	840
GTCAAGGGGA	TGTTTCTCAT	TTTCCATGAG	TGTCAGTTTT	CTTGCTATAT	TCCATGTCCT	900
ACAGTGACAT	TTCTAAATAT	TATACCTTTT	TCAGTTTTTT	TCACCATATT	TCACGTCCTA	960
AAGTATATAT	TTCTCATTTT	CCCTGATTTT	CAGTTTCCTT	GCCATATTCC	AGGTCCCTACA	1020
GTGTGCATTT	CTCATTTTTC	ACGTTTTTCA	GTAATTTCTT	CATTTTTTAA	GCCCTCAAAT	1080
GGATGTTTCT	CATTTTCCAT	GATTTTCAGT	TTTCTTGCCA	TATACCATGT	CCTACAGTGG	1140
ACATTTCTAA	ATTATCCACC	TTTTTCAGTT	TTTCATCGGC	ACATTTACAG	TCCTAAAGTG	1200
TGTATTTCTA	ATTTTCAGTG	ATTTTCAGTT	TTCTCGCCAT	ATTCCAGGAC	CTACAGTGTG	1260
CATTTCTCAT	TTTTTCACGT	TTTCAGTGAA	TTC			1293

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1044 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AGGCCTATGG	TGAAAAAGGA	AATATCTTCC	CCTGAAAAC	AGACAGAAGG	ATTCTCAGAA	60
TCTTATTTGT	GATGTGCGCC	CCTCAACTAA	CAGTGTTGAA	GCTTTCTTTT	GATAGAGCAG	120
TTTTGAAACA	CTCTTTTGT	AAAATCTGCA	AGAGGATATT	TGGATAGCTT	TGAGGATTTT	180
CGTTGGAAAC	GCGATTGTCT	TCATATAAAC	CCTAGACAGA	AGCATTCTCA	GAAGCTTCAT	240
TGGGATGTTT	CAGTTGAAGT	CACAGTGTG	AACAGTCCCC	TTTCATAGAG	CAGGTTTGAA	300
ACACTCTTTT	TTGTAGTATC	TGGAAGTGGA	CATTTGGAGC	GATCTCAGGA	CTGCGGTGAA	360
AAAGGAAATA	TCTTCCAATA	AAAGCTAGAT	AGAGGCAATG	TCAGAAACCT	TTTTCATGAT	420
GTATCTACTC	AGCTAACAGA	GTTGAACCTT	CCTTTGAGAG	AGCAGTTTGT	AAACACTCTT	480
TTTGTGGAAT	CTGCAAGTGG	ATATTGTCT	AGCTTTGAGG	ATTTTCGTTG	GAAACGGGAT	540
TACATATAAA	AAGCAGACAG	CAGCATTCCC	AGAACTTCT	TTGTGATGTT	TGCATTCAAG	600
TCACAGAGTT	GAACATTCCC	TTTCATAGAG	CAGGTTTGAA	ACACACTTTT	TGATGTATCT	660
GGATGTGGAC	ATTTGCAGCG	CTTTCAGGCC	TAAGGTGAAA	AGGAAATATC	TTCCCCTGAA	720
AACTAGACAG	AAGCATTCTC	AGAACTTAT	TTGTGATGTG	CGCCCTCAAC	TAACAGTGT	780
GAAGCTTTCT	TTTGATAGAG	GCAGTTTGA	AACACTCTTT	TGTGGAATCT	GCAAGTGGAT	840
ATTTGTCTAG	CTTTGAGGAT	TTCTTTGGAA	ACGGGATTAC	ATATAAAAAG	CAGACAGCAG	900
CATTCCCAGA	ATCTTGTTTG	TGATGTTTGC	ATTCAAGTCA	CAGAGTTGAA	CATTCCTTTT	960
CAGAGAGCAG	GTTTGAACAC	TCTTTTATA	GTATCTGGAT	GTGGACATTT	GGAGCGCTTT	1020
CAGGGGGGAT	CCTCTAGAAT	TCCT				1044

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2492 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTISENSE: NO
 (v) FRAGMENT TYPE:
 (vi) ORIGINAL SOURCE:
 (ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTGCAGCTGG	GGGTCTCCAA	TCAGGCAGGG	CCCCCTTACT	ACTCAGATGG	GGTGGCCGAG	60
TAGGGGAAGG	GGGTGCAGGC	TGCATGAGTG	GACACAGCTG	TAGGACTACC	TGGGGGCTGT	120
GGATCTATGG	GGGTGGGGAG	AAGCCCAGTG	ACAGTGCCTA	GAAGAGACAA	GGTGGCCTGA	180
GAGGGTCTGA	GGAACATAGA	GCTGGCCATG	TTGGGGCCAG	GTCTCAAGCA	GGAAGTGAGG	240
AATGGGACAG	GCTTGAGGAT	ACTCTACTCA	GTAGCCAGGA	TAGCAAGGAG	GGCTTGGGGT	300
TGCTATCCTG	GGGTTCAACC	CCCCAGGTTG	AAGGCCCTGG	GGGAGATGGT	CCCAGGACAT	360
ATTACAATGG	ACACAGGAGG	TTGGGACACC	TGGAGTCACC	AAACAAAACC	ATGCCAAGAG	420
AGACCATGAG	TAGGGGTGTC	CAGTCCAGCC	CTCTGACTGA	GCTGCATTGT	TCAAATCCAA	480
AGGGCCCCTG	CTGCCACCTA	GTGGCTGATG	GCATCCACAT	GACCCTGGGC	CACACGCGTT	540
TAGGGTCTCT	GTGAAGACCA	AGATCCTTGT	TACATTGAAC	GACTCCTAAA	TGAGCAGAGA	600
TTTCCACCTA	TTCGAAACAA	TCACATAAAA	TCCATCCTGG	AAAAAGCCTG	GGGGATGGCA	660
CTAAGGCTAG	GGATAGGGTG	GGATGAAGAT	TATAGTTACA	GTAAGGGGTT	TAGGGTTAGG	720
GATCAACGTT	GGTTAGGAGT	TAGGGATACA	GTAGGGTACC	GGTAGGGTTA	GGGGTTAGGG	780
TTAGGGGTTA	GGGTTAGGGT	TAGGGTTAGG	GTTAGGGTTA	GGGGTTAGGG	GTTAGGGTTA	840
GGGTTAGGTT	TTGGGGTGGC	GTATTTTGGT	CTTATACGCT	GTGTTCCACT	GGCAATGAAA	900
AGAGTTCTTG	TTTTTCCTTC	AGCAATTTGT	CATTTTTTAA	AGAGTTTAGC	AATTCTAACA	960
GATATAGACC	AGCTGTGCTA	TCTCATTGTG	GTTTTCAATT	GTAACCACAT	TGTGGTTTCA	1020
ATGTGTTTAC	TTGCCATCTG	TAGATCTTCT	TTGCGTGAGG	TGTCTGTTCA	GATGTGTGTG	1080
CATTTCTTGN	NTTTNGGCTG	TTAACTTAT	TGTTTAGTTT	TAATAATTTT	TTATATATTT	1140
GAAGACAAAT	CTTTCTCAGA	TGTGTATTTG	CAAATATTTT	TTCAATATGA	GGCTTGCTTT	1200
TGTCTCTAAC	AAGGTCTCTT	CAGAGATAAC	TTAAATATAA	GAAATCCACA	CTGTCACCTC	1260
TTTTGTGTAT	ATCTACCTTT	TGTGTCAATT	GTTAAATTC	ATTACCAAAC	CCAAAGGCAG	1320
ATAGCTTTTC	TTCTATTGTT	TCTTCTAGAA	ATTTGTATAG	TTTTTGCATT	TTAGTGTAA	1380
GATGATTTTG	AGTGATTATT	TGTGTAAGTT	GTAAAGTTTT	CGTCTATATC	CATATCATT	1440
CTTATGGTTT	CCAATTAATC	GTTCCCTCAC	TATTTTTGGG	AAAGACACAG	GATAGTGGGC	1500
TTTGTTAGAG	TAGATAGGTA	GCTAGACATG	AACAGGAGGG	GGCCTCCTGG	AAAAGGGAAA	1560
GTCTGGGAAG	GCTCACCTGG	AGGACCACCA	AAAATTCACA	TATTAGTAGC	ATCTCTAGTG	1620
CTGGAGTGGA	TGGGCACTTG	TCAATTGTGT	GTAGGAGGGA	AAAGAGGTCC	TATGCAGAAA	1680
GAAACTCCCT	AGAATCCCTC	TGAAGATGCC	CCAATCATTC	ACTCTGCAAT	AAAAATGTCA	1740
GAATATTGCT	AGCTACATGC	TGATAAGGNN	AAAGGGGACA	TTCTTAAGTG	AAACCTGGCA	1800
CCATAAGTAC	AGATTAGGGC	AGAGAAGGAC	ATTCAAAAGA	GGCAGGCGCA	GTAGGTACAA	1860
ACGTGATCGC	TGTCAGTGTG	CCTGGGATGG	CGGGAAGGAG	GCTGGTGCCA	GAGTGGATT	1920
GTATTGATCA	CCACACATAT	ACCTCAACCA	ACAGTGAGGA	GGTCCCACAA	GCCTAAGTGG	1980
GGCAAGTTGG	GGAGCTAAGG	CAGTAGCAGG	AAAACCAGAC	AAAGAAAACA	GGTGGAGACT	2040
TGAGACAGAG	GCAGGAATGT	GAAGAAATCC	AAAATAAAAT	TCCCTGCACA	GGACTCTTAG	2100
GCTGTTTAA	GCATCGCTCA	GTCCCACTCC	TCCCTATTTT	TCTACAATAA	ACTCTTTACA	2160
CTGTGTTTCT	TTTCAATGAA	GTTATCTGCC	ATCTTTGTAT	TGCTCTTGG	TGAAAATGTT	2220
TCTTCCAAGT	TAAACAAGAA	CTGGGACATC	AGCTCTCCCC	AGTAATAGCT	CCGTTTCAGT	2280
TTGAATTTAC	AGAACTGATG	GGCTTAATAA	CTGGCGCTCT	GACTTTAGTG	GTGCAGGAGG	2340
CCGTCACACC	GGGACCAAGA	GTGCCCTGCC	TAGTCCCCAT	CTGCCCCGAG	GTGGCGGCTG	2400
CCTCGACACT	GACAGCAATA	GGGTCCGGCA	GTGTCCCCAG	CTGCCAGCAG	GGGGCGTACG	2460
ACGACTACAC	TGTGAGCAAG	AGGGCCCCTGC	AG			2492

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 28 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO
 (iv) ANTISENSE: NO
 (v) FRAGMENT TYPE:
 (vi) ORIGINAL SOURCE:
 (ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GGGGAATTCA TTGGGATGTT TCAGTTGA

28

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTISENSE: NO
 (v) FRAGMENT TYPE:
 (vi) ORIGINAL SOURCE:
 (ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGAAAGTCCC CCCTAGGAGA TCTTAAGGA

29

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: RNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTISENSE: NO
 (v) FRAGMENT TYPE:
 (vi) ORIGINAL SOURCE:
 (ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

CCGCTTAATA CTCTGATGAG TCCGTGAGGA CGAAACGCTC TCGCACC

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
 (iii) HYPOTHETICAL: NO
 (iv) ANTISENSE: NO
 (v) FRAGMENT TYPE:
 (vi) ORIGINAL SOURCE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGATTTAAAT TAATTAAGCC CGGGC

25

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TAAATTTAAT TAATTCGGGC CCGTCGA

27

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 69 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(D) OTHER INFORMATION IL-2 signal sequence

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

ATG TAC AGG ATG CAA CTC CTG TCT TGC ATT GCA CTA AGT CTT GCA CTT
Met Tyr Arg Met Gln Leu Leu Ser Cys Ile Ala Leu Ser Leu Ala Leu

48

GTC ACA AAC AGT GCA CCT ACT
Val Thr Asn Ser Ala Pro Thr

69

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 945 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(A) NAME/KEY: Coding Sequence

(B) LOCATION: 1...942

(D) OTHER INFORMATION: Renilla Reinformis Luciferase

(x) PUBLICATION INFORMATION:

PATENT NO.: 5,418,155

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AGC	TTA	AAG	ATG	ACT	TCG	AAA	GTT	TAT	GAT	CCA	GAA	CAA	AGG	AAA	CGG	48
Ser	Leu	Lys	Met	Thr	Ser	Lys	Val	Tyr	Asp	Pro	Glu	Gln	Arg	Lys	Arg	
1				5					10					15		
ATG	ATA	ACT	GGT	CCG	CAG	TGG	TGG	GCC	AGA	TGT	AAA	CAA	ATG	AAT	GTT	96
Met	Ile	Thr	Gly	Pro	Gln	Trp	Trp	Ala	Arg	Cys	Lys	Gln	Met	Asn	Val	
			20					25					30			
CTT	GAT	TCA	TTT	ATT	AAT	TAT	TAT	GAT	TCA	GAA	AAA	CAT	GCA	GAA	AAT	144
Leu	Asp	Ser	Phe	Ile	Asn	Tyr	Tyr	Asp	Ser	Glu	Lys	His	Ala	Glu	Asn	
		35				40						45				
GCT	GTT	ATT	TTT	TTA	CAT	GGT	AAC	GCG	GCC	TCT	TCT	TAT	TTA	TGG	CGA	192
Ala	Val	Ile	Phe	Leu	His	Gly	Asn	Ala	Ala	Ser	Ser	Tyr	Leu	Trp	Arg	
	50					55					60					
CAT	GTT	GTG	CCA	CAT	ATT	GAG	CCA	GTA	GCG	CGG	TGT	ATT	ATA	CCA	GAT	240
His	Val	Val	Pro	His	Ile	Glu	Pro	Val	Ala	Arg	Cys	Ile	Ile	Pro	Asp	
65					70				75						80	
CTT	ATT	GGT	ATG	GGC	AAA	TCA	GGC	AAA	TCT	GGT	AAT	GGT	TCT	TAT	AGG	288
Leu	Ile	Gly	Met	Gly	Lys	Ser	Gly	Lys	Ser	Gly	Asn	Gly	Ser	Tyr	Arg	
				85					90					95		
TTA	CTT	GAT	CAT	TAC	AAA	TAT	CTT	ACT	GCA	TGG	TTG	AAC	TTC	TTA	ATT	336
Leu	Leu	Asp	His	Tyr	Lys	Tyr	Leu	Thr	Ala	Trp	Leu	Asn	Phe	Leu	Ile	
			100					105					110			
TAC	CAA	AGA	AGA	TCA	TTT	TTT	GTC	GGC	CAT	GAT	TGG	GGT	GCT	TGT	TTG	384
Tyr	Gln	Arg	Arg	Ser	Phe	Phe	Val	Gly	His	Asp	Trp	Gly	Ala	Cys	Leu	
		115					120					125				
GCA	TTT	CAT	TAT	AGC	TAT	GAG	CAT	CAA	GAT	AAG	ATC	AAA	GCA	ATA	GTT	432
Ala	Phe	His	Tyr	Ser	Tyr	Glu	His	Gln	Asp	Lys	Ile	Lys	Ala	Ile	Val	
	130					135					140					
CAC	GCT	GAA	AGT	GTA	GTA	GAT	GTG	ATT	GAA	TCA	TGG	GAT	GAA	TGG	CCT	480
His	Ala	Glu	Ser	Val	Val	Asp	Val	Ile	Glu	Ser	Trp	Asp	Glu	Trp	Pro	
145				150						155					160	
GAT	ATT	GAA	GAA	GAT	ATT	GCG	TTG	ATC	AAA	TCT	GAA	GAA	GGA	GAA	AAA	528
Asp	Ile	Glu	Glu	Asp	Ile	Ala	Leu	Ile	Lys	Ser	Glu	Glu	Gly	Glu	Lys	
				165					170					175		
ATG	GTT	TTG	GAG	AAT	AAC	TTC	TTC	GTG	GAA	ACC	ATG	TTG	CCA	TCA	AAA	576
Met	Val	Leu	Glu	Asn	Asn	Phe	Phe	Val	Glu	Thr	Met	Leu	Pro	Ser	Lys	
			180					185					190			
ATC	ATG	AGA	AAG	TTA	GAA	CCA	GAA	GAA	TTT	GCA	GCA	TAT	CTT	GAA	CCA	624
Ile	Met	Arg	Lys	Leu	Glu	Pro	Glu	Glu	Phe	Ala	Ala	Tyr	Leu	Glu	Pro	
		195					200					205				
TTC	AAA	GAG	AAA	GGT	GAA	GTT	CGT	CGT	CCA	ACA	TTA	TCA	TGG	CCT	CGT	672
Phe	Lys	Glu	Lys	Gly	Glu	Val	Arg	Arg	Pro	Thr	Leu	Ser	Trp	Pro	Arg	
	210					215					220					

GAA ATC CCG TTA GTA AAA GGT GGT AAA CCT GAC GTT GTA CAA ATT GTT	720
Glu Ile Pro Leu Val Lys Gly Gly Lys Pro Asp Val Val Gln Ile Val	
225 230 235 240	
AGG AAT TAT AAT GCT TAT CTA CGT GCA AGT GAT GAT TTA CCA AAA ATG	768
Arg Asn Tyr Asn Ala Tyr Leu Arg Ala Ser Asp Asp Leu Pro Lys Met	
245 250 255	
TTT ATT GAA TCG GAT CCA GGA TTC TTT TCC AAT GCT ATT GTT GAA GGC	816
Phe Ile Glu Ser Asp Pro Gly Phe Phe Ser Asn Ala Ile Val Glu Gly	
260 265 270	
GCC AAG AAG TTT CCT AAT ACT GAA TTT GTC AAA GTA AAA GGT CTT CAT	864
Ala Lys Lys Phe Pro Asn Thr Glu Phe Val Lys Val Lys Gly Leu His	
275 280 285	
TTT TCG CAA GAA GAT GCA CCT GAT GAA ATG GGA AAA TAT ATC AAA TCG	912
Phe Ser Gln Glu Asp Ala Pro Asp Glu Met Gly Lys Tyr Ile Lys Ser	
290 295 300	
TTC GTT GAG CGA GTT CTC AAA AAT GAA CAA TAA	945
Phe Val Glu Arg Val Leu Lys Asn Glu Gln	
305 310	

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

TTTGAATTC A TGTACAGGAT GCAACTCCTG 30

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) HYPOTHETICAL: NO

(iv) ANTISENSE: NO

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

(ix) FEATURE:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

TTTGAATTCA GTAGGTGCAC TGTTTGTAC 30